

Gdańsk, 24.6.2022

Recenzja rozprawy doktorskiej mgr. inż. Patryka Widulińskiego
pt. " Algorytmy detekcji infekcji programów komputerowych inspirowane biologicznymi
mechanizmami immunologicznymi "

Zawartość rozprawy mieści się w szeroko pojętej tematyce metod obliczeniowych wspomagających ochronę bezpieczeństwa informacji w systemach komputerowych przy wykorzystaniu algorytmów inspirowanych biologicznie. Jest to tematyka bardzo aktualna zarówno ze względów merytorycznych, związanych z rosnącą skalą zagrożeń i skutków naruszenia bezpieczeństwa informacji, jak i metodologicznych, związanych z koniecznością przetwarzania ogromnych maszyn danych, gdzie zawodzą klasyczne modele analityczne oraz algorytmy klasyfikacji i optymalizacji. Wymierną tego oznaką jest lawinowy coroczny przyrost publikacji i rozwój narzędzi obliczeniowych opartych na algorytmach ewolucyjnych, genetycznych, sztucznych sieciach neuronowych, modelach rojów pszczół czy kolonii mrówek. Z punktu widzenia ochrony systemów informatycznych przed różnego rodzaju intruzami od kilku lat duże nadzieje budzą mechanizmy wywołujące efekty podobne do reakcji immunologicznej w organizmach żywych w obliczu zewnętrznej infekcji. Doprowadziło to do wyodrębnienia ważnego obszaru badań nad sztucznymi systemami immunologicznymi, zaś kluczowe pojęcia biologiczne, takie jak organizm, komórka, receptor, limfocyt, antygen, infekcja, czy odporność swoista i nieswoista, szybko znalazły swe cyfrowe odpowiedniki i modele. Na tle tego stanu wiedzy Doktorant zaproponował usprawnienie metod ochrony ustalonego organizmu statycznego przed pewną klasą infekcji - mianowicie ochrony programu przechowywanego w pamięci lokalnej systemu komputerowego przed nieuprawnioną modyfikacją jego kodu binarnego. Jest to więc problem zapewniania integralności danych; w rozważanej w rozprawie wersji ma on znacznie szersze zastosowania, np. transmisja danych w obecności szumów i zakłóceń, spójność w rozproszonych bazach danych, czy też dystrybucja treści w sieciach informacyjnych, a ostatnio zwłaszcza kontrola autentyczności i integralności przekazów w mediach społecznościowych. Znane od dawna podejścia obejmują m.in. wykorzystanie kodów nadmiarowych oraz kryptograficznych funkcji skrótu, jak również zaawansowane protokoły głosowania zgodnie z zasadą *lots of copies keep stuff safe*. Naturalnie najprostszym sposobem byłoby porównywanie kodu z jego bezpiecznie przechowywaną oryginalną kopią, co byłoby jednak pamięciowo nieefektywne; dla oszczędności pamięci można ograniczyć się do weryfikacji odpowiednich sum kontrolnych lub wartości skrótu, co z kolei pogarsza wykrywalność i lokalizację intruzji. Sztuczne mechanizmy immunologiczne z odpowiednio wytrenowanym zbiorem receptorów przy zastosowaniu negatywnej selekcji potencjalnie dostarczają tu pożądanego kompromisu, dlatego uważam, że dokonany przez Doktoranta wybór takiego właśnie podejścia jest w pełni uzasadniony.

Postawione w rozprawie zagadnienie trenowania receptorów dla danego organizmu chronionego, tj. zestawu danych trenujących, można streścić następująco (co Autor czyni jedynie w sposób opisowy – jak się wydaje, ze szkodą dla przejrzystości swoich wywodów): mając l -wymiarową przestrzeń Hamminga Z_l , gdzie l jest zadany parametrem określającym rozmiar komórki organizmu, tj. ciągu własnego, oraz zadany zbiór ciągów własnych $S \subseteq Z_l$, należy skonstruować zbiór ciągów-receptorów $R \subseteq Z_l$ spełniający warunek "rozpoznawalności swój-obcy". Dla danego parametru $m \leq l$, zwanego progiem aktywacji, i dla dowolnego ciągu $\mathbf{a} \in Z_l$ zdefiniujmy $\mathbf{a}^{(k,m)}$ jako m -elementowy podciąg ciągu $\mathbf{a}$ rozpo-

czynający się od pozycji k . Pierwszą część warunku ("swój") można sformułować w następującej postaci:

$$\forall s \in S, r \in R, k \in \{1, \dots, l-m+1\}: r^{(k,m)} \neq s^{(k,m)}.$$

Oznacza to, że żaden receptor nie zostanie aktywowany przy zestawieniu z jednym z ciągów własnych. Drugą część warunku ("obcy") dla danego zbioru ciągów obcych $T \subseteq Z_l \setminus S$ można sformułować w następującej postaci:

$$\forall t \in T: \exists r \in R, k \in \{1, \dots, l-m+1\}: r^{(k,m)} = t^{(k,m)}.$$

Oznacza to, że każdy ciąg obcy aktywuje przynajmniej jeden receptor. Ze względu na kształtowanie właściwej relacji między skutecznością ochrony a jej kosztami postuluje się, by warunek "rozpoznawalności swój-obcy" był spełniony dla jak najmniejszego zbioru R i jak największego zbioru T . Naturalnie powyższa szkicowa formalizacja jest tylko jedną z możliwych, a przytoczona jest tu jako najbardziej zbliżona do ujęcia przedstawionego w rozprawie. Może ona zostać uznana za problem czysto kombinatoryczny, niewymagający odniesień do aparatu pojęciowego sztucznych systemów immunologicznych. Trudno powiedzieć, do jakiego stopnia aparat ten okazał się niezbędny Doktorantowi przy pisaniu pierwszej części rozdziału 3, stanowiącego teoretyczną podbudowę proponowanego rozwiązania – być może zadecydowała wygoda posługiwania się ustaloną terminologią biologiczną i zestawem uznanych algorytmów generacji receptorów.

Pozostając konsekwentnie w obrębie algorytmów negatywnej selekcji, po rozważeniu algorytmów losowych oraz algorytmów opartych na wzorcach wewnątrzkomórkowych Autor słusznie ocenił te ostatnie jako bardziej przydatne. Dalsza część rozważań w rozdziale 3 poświęcona jest generacji receptorów na podstawie wzorców przy pomocy tablic filtrujących i drzew binarnych. Nie są to metody oryginalne – Autor opierał się tu głównie na wcześniejszych pracach [77] i [78] – stąd ich tak szeroki opis nieco zaburza strukturę rozprawy i utrudnia identyfikację pierwiastka oryginalnego. Jednak dzięki takiemu układowi treści Autorowi udało się przekonująco zarysować podstawowy problem klasycznych algorytmów selekcji – przywiązanie do sztywnego podziału chronionego organizmu na komórki (tj. kodu programu na ciągi własne), i jego proponowane rozwiązanie, tj. wprowadzenie dodatkowego zbioru receptorów międzykomórkowych. Jak wskazuje nazwa, są one aktywowane przez anomalie (tj. modyfikacje kodu) obejmujące krótką ostatnią część poprzedniego i krótką początkową część następnego ciągu własnego, niemożliwe do wykrycia przez receptory podstawowe. Trafnie zakreśla się przy tym obszar możliwych zastosowań do organizmów statycznych w przeciwieństwie np. do ruchu sieciowego, gdzie dynamika środowiska i relacje czasowe (*timing*) narzucone przez protokoły komunikacyjne bardzo by skomplikowały model zachowań organizmu, tym samym także wykrywanych anomalii. Sama metoda generacji receptorów międzykomórkowych nie odbiega od metod przedstawionych uprzednio, toteż jej znaczenie teoretyczne jest niewielkie, zaś jej głównych zalet należy upatrywać w omawianych dalej wynikach eksperymentalnych, wskazujących na możliwość polepszenia przy jej pomocy wymienności wskaźników jakości ochrony kodu przed intruzją.

Na podstawie lektury rozdziału 3 rozprawy oraz biorąc pod uwagę wspomniane wyniki eksperymentów, obszernie zrelacjonowane i przeanalizowane w rozdziale 4 rozprawy, oceniam, że Doktorant rozwiązał sformułowane we wprowadzeniu do rozprawy zagadnienie ochrony kodu programu w lokalnej pamięci systemu komputerowego przed rozważanym typem intruzji. W szczególności w punkcie 4.1 rozprawy zademonstrował, że w obecności losowo wprowadzanych intruzji uzupełnienie algorytmu negatywnej selekcji o generację receptorów międzykomórkowych może polepszyć wymiennosc pomiędzy głównymi wskaźnikami jakości, to jest: stopniem wykrywalności intruzji, szybkością działania algorytmu i zajętością pamięci przez zbiór wygenerowanych receptorów relatywnie do objętości oryginalnego kodu. W punkcie 4.1 rozprawy dokonane zostało porównanie metody receptorów międzykomórkowych z wybranymi istniejącymi algorytmami negatywnej selekcji, z którego wychodzi ona obronną ręką zwłaszcza w warunkach redukcji zbioru danych trenujących. Opracowanie metody opartej na receptorach międzykomórkowych uważam za główny wkład rozprawy do aktualnego stanu wiedzy o sztucznych systemach odpornościowych, obok poznawczej wartości uzyskanych wyników eksperymentalnych. Wkład ten mógłby być bardziej znaczący, gdyby badania oparto na analizie większej populacji programów lub innych zbiorów trenujących, oraz gdyby podjęto przynajmniej próbę uzyskania bardziej uniwersalnych wyników w drodze przybliżonej analizy probabilistycznej. Rozprawa zyskałaby także więcej naukowych i praktycznych walorów, gdyby repertuar rozważanych intruzji rozszerzyć

np. o lokalizację wyraźnie skorelowane z kodem programu, bądź anomalie ujawniające się dopiero w trakcie wykonania programu.

Podsumowując zawartość merytoryczną rozprawy należy ocenić, że do rozwiązania postawionego zagadnienia Doktorant użył właściwych metod, wykazując się znacznymi kompetencjami w zakresie sztucznych mechanizmów immunologicznych oraz świadomością rozmaitych zagrożeń bezpieczeństwa informacji we współczesnych systemach komputerowych. Świadczy o tym pośrednio spis cytowanej literatury, obejmujący 90 pozycji. Są wśród nich pozycje autorskie bądź współautorskie w liczbie siedem, w tym cztery prezentowane na konferencjach międzynarodowych o mniejszej randze i jedna opublikowana w czasopiśmie z listy MNiSW, "wartym" 70 punktów. W tej obszernej bibliografii zwraca jednak uwagę brak kilku znaczących pozycji przeglądowych kładących nacisk na zastosowania w systemach IDS, jak np. prace Fernandes i in. z 2017 r. (DOI 10.1016/j.jisa.2017.06.007) oraz Hua Yang i in. z 2014 r. (DOI 10.1155/2014/156790), a także w optymalizacji wielokryterialnej, jak np. praca Lingjie Li i in. z 2022 r. (DOI j.neucom.2021.08.154). Zebraną literaturę warto by ogólnie "odświeżyć", pewnym mankamentem jest bowiem fakt, że ponad połowa cytowanych pozycji pochodzi sprzed 10 i więcej lat. Wreszcie można odnieść wrażenie, że wiele frapujących zagadnień związanych ze sztucznymi odpowiedziami immunologicznymi na intruzje w systemach komputerowych, zarówno od strony algorytmów optymalnej generacji receptorów, jak i ich analizy statystycznej oraz repertuaru możliwych anomalii, znalazła dość nisko odbicie w oryginalnych fragmentach rozprawy, obejmujących rozdziały 3 i 4.

Od strony prezentacyjnej i redakcyjnej rozprawa generalnie sytuuje się na poziomie charakterystycznym dla tego typu dokumentów, można jednak wysunąć kilka zastrzeżeń różnego rodzaju.

- Brak matematycznego formalizmu wydłuża tekst, zmusza czytelnika, by domyślał się istoty wielu przedstawianych zagadnień i utrudnia ocenę ich stopnia trudności. Powoduje też wrażenie braku jednolitej organizacji przeglądu literatury i umiejscowienia własnych propozycji w odpowiednim kontekście. Opisowy styl rozprawy prowadzi do wielu niejasności, np. stwierdzenia typu "liczby receptorów potrzebnych do pokrycia przestrzeni ciągów obcych" (s. 14) bądź "szacowania zasięgu receptorów w NSA" (s. 15) bez dodatkowego objaśnienia można rozumieć na wiele sposobów. Ta sama uwaga dotyczy używanego w przeglądzie literatury pojęcia "dziur", które przecież łatwo by-łoby zdefiniować przy pomocy wyżej naszkicowanej formalizacji.
- Niekiedy operuje się pojęciami uprzednio niewprowadzonymi, np. stwierdzenie na s. 9 "działanie polega na generacji ciągów binarnych, które potrafią dopasowywać ciągi obce, a jednocześnie nie dopasowują nigdy ciągów własnych" staje się bardziej zrozumiałe dopiero po spojrzeniu na rys. 12 na s. 32. Stosuje się odwołania w przód, np. w nawiązaniu do wariantowych kryteriów TP_i i FN_i lub receptorów podstawowych (ss. 12, 21).
- Niektóre oznaczenia są mylące (np. litera R stosowana jest zarówno dla liczby, jak i zbioru receptorów, a także dla zajętości pamięci), inne są niekonsekwentne (np. moc zbioru oznaczana jest jako $num(\cdot)$ innym razem jako $|\cdot|$, podobnie do wartości bezwzględnej). Kryteria oceny algorytmu zapisywane są w różnych miejscach różnym krojem czcionki. Bardzo by się przydał spis oznaczeń.
- Wyniki eksperymentalne w punkcie 4.1 rozprawy powinny moim zdaniem zostać przedstawione bardziej profesjonalnie, naturalna byłaby tu konwencja optymalizacji kompromisowej, np. w postaci powierzchni Pareto. Autor próbował znaleźć receptę na przejrzyste uporządkowanie dużej masy danych wynikowych poprzez jednowymiarowe rankingowanie, ale chyba bez sukcesu, przez co wnioski wydają się nieco incydentalne. Nie są też wsparte analizą wiarygodności statystycznej.
- Niektóre fragmenty rozprawy wydają się banalne lub nadmiarowe, jak np. rys. 5 i 9 oraz rozumowanie probabilistyczne na s. 39. Natomiast całość analizy probabilistycznej w punkcie 3.6 rozprawy nie wydaje się być przydatna w dalszych rozważaniach.
- Nieprecyzyjny język wywodu często utrudnia lekturę; ilustrują to np. sformułowania: "w związku z naturą podmiotowych ciągów własnych, do dalszych rozważań w niniejszej rozprawie wybrany został algorytm selekcji negatywnej" (s. 9 – nie wiadomo, o jaką naturę chodzi); "Dopasowanie polega na wystąpieniu zarówno w receptorze, jak i w sprawdzanym ciągu binarnym m identycznych kolejnych bitów na jednakowej pozycji k " (s. 10); "NSA zakłada istnienie jednego zbioru receptorów R , który zawiera wszystkie wygenerowane receptory" (s. 11 – skoro zostały wygenerowane, to ich zbiór istnieje); "Receptory wygenerowane przez NSA z reguły nie pozwalają na uzyskanie 100% wykrywalności anomalii" (s. 11 – nie jest wyjaśnione, dlaczego); "ciągi binarne receptorów powstają przez odpytanie generatora liczb pseudolosowych i konkatencję bajtów dopóki nie powstanie R_{max} ciągów o długości co najmniej l " (s. 19); "generacja polega na utworzeniu tablicy wzorców dla każdej kombinacji bitów progu aktywacji" (s. 19); "Algorytm negatywnej selekcji wyko-

rzystywany jest do wykrywania anomalii typowo w ruchu sieciowym między innymi ze względu na strumieniowy charakter danych do skanowania" (s. 20), "instrukcje wykonywane są przez procesor kolejno, lub jedna po drugiej" (s. 20), "możliwe jest wykrycie infekcji, o których sygnaturach nie wiedział wcześniej IDS" (s. 20), "anomalie nie mogą wystąpić między pakietami sieciowymi" (s. 35) i wiele innych.

- Styl stosowany zwłaszcza w odniesieniu do omawianych wyników numerycznych jest niekiedy rażąco przesadny. Doktorant wielokrotnie używa przymiotnika "znakomity", a niekiedy też "imponujący" lub "atrakcyjny"; kiedy indziej bez głębszego uzasadnienia określa wyniki lub ich zastosowania jako "nieopłacalne", "nieadekwatne", bądź "nierелеwantne".

Wymienione usterki i zarzuty obniżają ogólną ocenę rozprawy, jednak nie przekreślają wartości poznawczej opracowanych metod i uzyskanych wyników. Nie podważają też wykazanych przez Doktoranta kompetencji merytorycznych i metodologicznej poprawności. W świetle tego stwierdzam, że rozprawa doktorska mgr. inż. Patryka Widulińskiego spełnia wymagania obowiązującej ustawy o stopniach naukowych i tytule naukowym i wnoszę o jej dopuszczenie do publicznej obrony.

